

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-337981
 (43)Date of publication of application : 07.12.2001

(51)Int.Cl.

G06F 17/30
 G06F 17/28

(21)Application number : 2000-159702
 (22)Date of filing : 30.05.2000

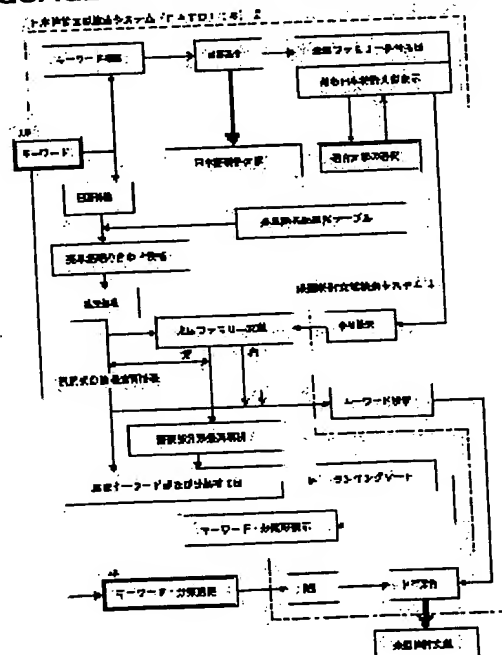
(71)Applicant : PATOLIS CORP
 (72)Inventor : NAKATANI YOICHI
 TANAKA MASAO
 HIGUCHI SHIGETO
 FUKUI MASATOSHI

(54) RETRIEVAL EXPRESSION CONVERTING METHOD FOR CONVERTING FIRST LANGUAGE RETRIEVAL EXPRESSION TO SECOND LANGUAGE RETRIEVAL EXPRESSION

(57)Abstract:

PROBLEM TO BE SOLVED: To effectively use the result obtained by retrieving a patent document database constructed in Japanese and a Japanese retrieval expression used therefor for the formation of an English retrieval expression for patent document database retrieval constructed in English or to make available a US patent document without particularly forming the English retrieval expression by actively use the patent family relation of patent documents.

SOLUTION: A retrieval is executed on a Japanese Patent document retrieval system 2 by use of the Japanese retrieval expression, and a family document number having a patent family in US is extracted from the resulting Japanese retrieval expression. Each keyword of the Japanese retrieval expression is translated to an English word (generally, a plurality of words), the keyword of the Japanese retrieval expression is substituted by a translated English word such that the keyword (English word) of the English word imparted to the family document is conformed thereto, whereby the English retrieval expression is formed to execute a retrieval in a US patent document retrieval system 3.



LEGAL STATUS

[Date of request for examination]
 [Date of sending the examiner's decision of rejection]
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
 [Date of final disposal for application]
 [Patent number]

BEST AVAILABLE COPY

【特許請求の範囲】

【請求項 1】 第 1 言語のキーワードを用いた第 1 言語検索式を作成する第 1 言語検索式作成工程と、

上記第 1 言語検索式を用いて、第 1 の言語に基づいて構築された第 1 のキーワード特許文献検索システム上で検索を実行する第 1 検索実行工程と、

上記第 1 検索実行工程によって得られた第 1 言語の特許文献の中で第 2 言語の特許文献をパテントファミリーとして有している第 2 言語のファミリー文献の番号を抽出するファミリー文献番号抽出工程と、

上記第 1 言語検索式のキーワードのそれぞれを単数又は複数の第 2 の言語の単語に対訳する対訳工程と、

上記対訳工程において対訳された第 2 言語の単数又は複数の単語の中から、上記ファミリー文献番号抽出工程で抽出された番号を有する第 2 言語のファミリー文献に付与されている第 2 言語キーワードに一致する第 2 言語の単語を選択する単語選択工程と、

上記単語選択工程において選択された第 2 言語の単語を上記第 1 言語検索式のキーワードと置き換えて第 2 言語検索式とする置き換え工程と、からなることを特徴とする第 1 言語検索式を第 2 言語検索式に変換するための検索式変換方法。

【請求項 2】 請求項 1 に記載された検索式変換方法をコンピュータに実行させるためのプログラムが記録された記録媒体。

【請求項 3】 第 1 の言語に基づいて構築された第 1 のキーワード特許文献検索システムと第 2 の言語に基づいて構築された第 2 のキーワード特許文献検索システムとの間に介在して両者との間で通信可能に接続されるための複数言語間の特許文献検索のためのインターフェースシステムであって、この複数言語間の特許文献検索のためのインターフェースシステムは、

上記第 1 のキーワード特許文献検索システムのための第 1 言語検索式を入力するための第 1 検索式入力手段と、上記第 1 言語検索式により上記第 1 のキーワード特許文献検索システムにおいて検索されることにより適合した特許文献にかかる特許文献情報に基づいて、上記適合した特許文献の中から上記第 2 の言語によって記載されたパテントファミリー文献番号を抽出するファミリー番号抽出手段と、

上記第 1 言語検索式のキーワードのそれぞれを単数又は複数の第 2 の言語の単語に対訳する対訳手段と、

上記対訳手段によって対訳された第 2 言語の単数又は複数の単語の中から、上記ファミリー番号抽出手段で抽出された番号を有する第 2 言語のファミリー文献に付与されている第 2 言語キーワードに一致する第 2 言語の単語を選択する単語選択手段と、

上記単語選択手段によって選択された第 2 言語の単語を上記第 1 言語検索式のキーワードと置き換えて第 2 言語検索式とする置き換え手段と、を備えており、

上記第 2 言語検索式により、上記第 2 の言語によって記載された特許文献を対象にして、第 2 のキーワード特許文献検索システムに、キーワード検索を実行させるものであることを特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【請求項 4】 請求項 3 に記載された複数言語間の特許文献検索のためのインターフェースシステムにおいて、このインターフェースシステムは、更に、

上記置き換え手段によって、上記第 1 言語検索式のキーワードの全てが、上記第 2 言語のファミリー文献に付与されている第 2 言語キーワードに置き換えられなかったとき、第 2 の言語の特許文献の文面又は第 2 の言語のキーワードを表示するための第 2 言語文献表示手段と、

上記第 2 言語文献表示手段によって表示された、第 2 の言語の特許文献の文面又は第 2 の言語のキーワード、から選んで重要なキーワードを入力又は選択するためのキーワード選択手段と、を備えており、

上記キーワード選択手段によって入力又は選択されたキーワードにより、上記第 2 の言語によって記載された特許文献を対象にして、第 2 のキーワード特許文献検索システムに、キーワード検索を実行させるものであることを特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【請求項 5】 請求項 3 又は請求項 4 のいずれかに記載された複数言語間の特許文献検索のためのインターフェースシステムにおいて、このインターフェースシステムは、更に、

上記第 2 の言語によって記載された特許文献を対象にして、第 2 のキーワード特許文献検索システムによって実行されたキーワード検索の結果を出力するための出力手段を備えること、を特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【請求項 6】 上記請求項 3 から請求項 5 までのいずれかに記載された複数言語間の特許文献検索のためのインターフェースシステムをコンピュータシステムに実現させるためのプログラムが記録された記録媒体。

【請求項 7】 それぞれが異なる言語で構築された複数の文献検索システムを含み、同一の内容の文献が上記各文献検索システムのデータベースに散在している複数言語検索システムのための検索方法であって、この検索方法は、

一つの言語により作成された質問を上記複数の言語のうちのこの言語以外の言語に対応する翻訳候補を作成するための対訳辞書と、上記一つの言語により作成された質問に対する回答文献中の一部に存在し、各言語により記述された同一の内容の文献とを用いて、上記翻訳候補の中から当該質問に対応する各言語の質問を選定すること、を特徴とする複数言語検索システムのための検索方法。

【請求項 8】 それぞれが異なる言語で構築された複数の文献検索システムを含み、同一の内容の文献が上記各

文献検索システムのデータベースに散在している複数言語検索システムであって、この複数言語検索システムは、

一つの言語により作成された質問を上記複数の言語のうちのこの言語以外の言語に対応する翻訳候補を作成するための対訳辞書と、上記一つの言語により作成された質問に対する回答文献中の一部に存在し、各言語により記述された同一の内容の文献とを用いて、上記翻訳候補の中から当該質問に対応する各言語の質問を選定するための質問言語変換手段を備えていることを特徴とする複数言語検索システム。

【請求項 9】 それぞれが異なる言語で構築された複数の文献検索システムを含み、同一の内容の文献が上記各文献検索システムのデータベースに散在している複数言語検索システムのためのコンピュータプログラムが記録された記録媒体であって、この記録媒体に記録されたコンピュータプログラムは、

一つの言語により作成された質問を上記複数の言語のうちのこの言語以外の言語に対応する翻訳候補を作成するための対訳辞書と、上記一つの言語により作成された質問に対する回答文献中の一部に存在し、各言語により記述された同一の内容の文献とを用いて、上記翻訳候補の中から当該質問に対応する各言語の質問を選定するための質問言語変換機能を上記複数言語検索システムに実現させるコンピュータプログラムであることを特徴とする記録媒体。

【請求項 10】 第 1 の言語に基づいて構築された第 1 のキーワード特許文献検索システムと、第 2 の言語に基づいて構築された第 2 のキーワード特許文献検索システムと、

上記 2 つのキーワード特許文献検索システムの間を介して両者との間で通信可能に接続されるための複数言語間の特許文献検索のためのインターフェースシステムとを備えた複数言語間の特許文献検索システムであって、上記インターフェースシステムは、

上記第 1 のキーワード特許文献検索システムのための第 1 言語検索式を入力するための第 1 検索式入力手段と、上記第 1 言語検索式により上記第 1 のキーワード特許文献検索システムにおいて検索されることにより適合した特許文献にかかる特許文献情報に基づいて、上記適合した特許文献の中から上記第 2 の言語によって記載されたパテントファミリー文献番号を抽出するファミリー番号抽出手段と、

上記第 1 言語検索式のキーワードのそれぞれを単数又は複数の第 2 の言語の単語に対訳する対訳手段と、

上記対訳手段によって対訳された第 2 言語の単数又は複数の単語の中から、上記ファミリー番号抽出手段で抽出された番号を有する第 2 言語のファミリー文献に付与されている第 2 言語キーワードに一致する第 2 言語の単語を選択する単語選択手段と、

上記単語選択手段によって選択された第 2 言語の単語を上記第 1 言語検索式のキーワードと置き換えて第 2 言語検索式とする置き換え手段と、を備えており、

上記第 2 言語検索式により、上記第 2 の言語によって記載された特許文献を対象にして、第 2 のキーワード特許文献検索システムに、キーワード検索を実行させるインターフェースシステムであることを特徴とする複数言語間の特許文献検索システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は複数言語間の検索システム、特に特許文献検索に適した検索システムに関する。更に詳しくは、第 1 言語検索式を第 2 言語検索式に変換するための検索式変換方法、複数言語間の特許文献検索のためのインターフェースシステム、複数言語検索システムのための検索方法に関する。

【0002】

【従来の技術】文献検索をするためデータベースにアクセスしようとするとき、機械検索では検索式といわれる質問文あるいは質問式が作成される。この検索式によって、得たいと希望する文献を特定するための種々の条件が指定される。

【0003】一般にデータベースは言語に依存する。つまり、ある国で使用されているデータベースはその国の言語を基礎に構築されている。これは特許文献のデータベースにおいても同様である。したがって、例えば、日本人が、米国特許文献を入手しようとしたとき、英語によって構築されている米国特許文献のデータベースにアクセスできるようにするために、英語の検索式を作成しなければならない。

【0004】このとき、英語の検索式を得るためには、日本人が自らの力によって当初から英語の検索式を作成する方法、及び、日本語で作成した検索式を、例えば自動翻訳システムを用いて、翻訳する方法が考えられる。

【0005】機械検索には、技術的な知識、同義語、関連語を含む英語に関する知識、及び、使用するデータベース（検索システム）の特性に関する知識等が同時に広く必要とされるので、多くの日本人にとっては、いきなり英語の検索式を作成するという前者の方法はなかなか採りがたく、これをしたとしても結果として信頼のできる検索回答を得ることができないという問題がある。

【0006】また、考えられる後者の方法では、日本語の単語の意味と英語の単語の意味の間には必ずしも一対一の対応がとれないことから、更に、自動翻訳システムが検索式以外に参照すべき検索しようとするテーマの情報を持たないことから、自動翻訳システムが対訳辞書にある複数の英単語の中から適切な英単語を選択できるようなことがきわめて困難であるのが実状である。

【0007】

【発明が解決しようとする課題】ところで、パリ同盟条

約等の優先権を主張した出願、及び、国際特許協力条約に基づく国際出願等によって、世界各国には、実質的にそれぞれの国の言語に翻訳され、実質的に同一の技術内容が記載された特許文献が数多く存在している。

【0008】このような関係にある複数の国の特許文献はパテントファミリーといわれている。従来、このようにパテントファミリー関係にある特許文献が数多く存在しているにもかかわらず、いわゆる「ファミリー検索」といわれる検索法、つまり、予めわかっている特許文献に記載された第1国の出願番号（優先権番号）や国際出願番号をタグとして利用して第2国の特許文献にアクセスするという検索法以外には、この事実が、特許文献の検索、あるいは検索式の作成のために積極的に利用されることはなかった。

【0009】パテントファミリー関係にある2つの文献には、実質的に同じ内容が記載されているのであるから、ある技術に着目したとき、この技術を記載した日本語の特許文献とファミリー関係にある米国特許文献には、この技術を表現するために必要な英語の単語、これらの組み合わせ、表現方法のヒントが必ずや含まれているはずである。それ故、外国語を基礎に構築されたキーワード特許文献検索システムにおいて入力しなければならない外国語の検索式を作成するためには、このことが大いに役に立つはずである。

【0010】第1国への特許出願（比喩的にいえばマザーパテント）が日本語特許文献、及び、これを優先権主張の根拠として出願された第2国へ特許出願（比喩的にいえばドーターパテント）が米国特許文献であるとき、この米国特許文献は、英語国人によって書かれたものであるから、日本語検索式のキーワードを英語検索式に変換するときの手がかりを与えるものと考えられる。

【0011】本発明は、このような特許文献のパテントファミリー関係を積極的に利用し、これにより日本語で構築された特許文献データベースを検索して得られた成果、及び、このときの検索式を英語で構築された特許文献データベースの検索に有効に利用しようとするのが課題とされるものである。

【0012】更に本発明は、日本語検索式を使用して日本語特許文献を入手すると同時に、特に英語検索式を人が作成することなく、米国特許文献も入手できるようにすることを課題とするものである。

【0013】

【課題を解決するための手段】本発明では、日本語の特許文献検索システムにおいて日本語キーワードからなる検索式によって検索が行われる。この検索式によってヒットした日本の特許文献の中から上記パテントファミリー関係にある米国の特許文献が抽出される。

【0014】一方、日英対訳辞書により上記日本語キーワードのそれぞれに対応する英語単語が抽出される。多くの場合、一つの日本語キーワードに対して複数の英単

語が対応する。

【0015】そのため、仮に、抽出された英単語の全てを含むように単純に論理和で結合して、これを上記検索式の日本語キーワードと置き換えて英語検索式とした場合、この英語検索式を用いた検索によってヒットする米国特許文献の数が多いに多くなり過ぎ、実質的に検索の意味をなさないであろう。

【0016】そこで、一つ一つの日本語キーワードにそれぞれ対応する複数の英単語群の中から、統計的手法により順位付けが行われる。たとえば、英単語の共起頻度テーブルを参照しながら、英単語の組み合わせ候補群が選出される。

【0017】更に、この英単語の組み合わせ候補群の中から、上述のパテントファミリー文献（米国特許文献、抽出された日本語特許文献と対応する米国の特許文献）に付与されている英語キーワード（通常は複数）に一致する英単語が選択されて、日本語検索式の日本語キーワードのそれぞれに置き換えられる。

【0018】こうして、最初の日本語検索式は、その式の形（論理式構造）を保ったまま、日本語キーワードが英語キーワードに置き換えられることにより、英語検索式が作成されたことになる。そして、次は米国特許文献検索システムにおいて、この英語検索式を用いて検索が実行される。

【0019】以上では、日本語検索式が人手により作成される以外は、全て自動的に実行される。このような自動実行により、全ての日本語キーワードが英語キーワードに置き換えられないとき、又はこの検索式により満足できる回答が得られないとき、例えば、技術分野の不適合、最初の日本語検索式の不出来、統計的手法による処理の技術範囲への不適合、参照される米国特許文献に付与された英語キーワードの不備、その他様々の要因によって、満足できる回答が得られないことが起こりうる。

【0020】このとき、人手によって英語の検索式を作成する必要があるが、この場合でも、本発明では、パテントファミリーである米国特許文献とこれが持つデータが利用される。特に米国特許文献がマザーパテントであるとき、米国人によって作成、あるいは選出された特許明細書、アブストラクト、キーワードは、日本人が英語キーワードを選択する上で大いに参考になるので、これら及び特許分類を参照しながら英語検索式を作成することができる。

【0021】本発明は、また、既存の、あるいは将来開発される日本語の特許文献検索システム及び英語の特許文献検索システムが利用される。そして一つの発明はこれら2つのシステムの間に介在することによって英語の特許文献検索システムを日本人がより利用しやすくするものである。それ故、本発明は、一種のインターフェース、あるいは、英語特許文献検索システムのための支援システムと考えることができる。また、あるいは、ユー

ザー（主人）の命を受けて、2つの特許文献検索システムの間を情報を持って行き来することにより、サービスを行うようにみえることから、ロボット（検索支援ロボット）に例えることができる。

【0022】更に、上記したような、日本語と英語という2つの言語の関係は、逆の場合、すなわち、英語と日本語との関係であっても適用でき、また、各国で 사용되는複数の他の言語間の関係（英語－中国語、英語－ドイツ語、その他）であって適用可能である。それゆえ、本発明を利用することにより、異なる言語で記載された世界中の特許文献へのアクセスが容易に行うことができるようになる。

【0023】更に、ユーザーが、インターネット、パソコン通信、LAN等の、ネットワーク、そして、自前又は商用サービスを問わず任意の通信手段を介して行われるところの特許文献検索サービスを利用するときにも、本発明を利用することができるので、この点から本発明は特許文献検索のための支援システムとも考えることができる。

【0024】更に、本発明は、日本語の特許文献検索システム及び英語の特許文献検索システムをサブシステムとして持つところの、一つの大きな複数言語の特許文献のための機械検索システムとして考えることができる。

【0025】そして、上記課題は以下の解決手段によって解決される。すなわち、

「第1番目の発明の解決手段」第1言語のキーワードを用いた第1言語検索式を作成する第1言語検索式作成工程と、上記第1言語検索式を用いて、第1の言語に基づいて構築された第1のキーワード特許文献検索システム上で検索を実行する第1検索実行工程と、上記第1検索実行工程によって得られた第1言語の特許文献の中で第2言語の特許文献をパテントファミリーとして有している第2言語のファミリー文献の番号を抽出するファミリー文献番号抽出工程と、上記第1言語検索式のキーワードのそれぞれを単数又は複数の第2の言語の単語に対訳する対訳工程と、上記対訳工程において対訳された第2言語の単数又は複数の単語の中から、上記ファミリー文献番号抽出工程で抽出された番号を有する第2言語のファミリー文献に付与されている第2言語キーワードに一致する第2言語の単語を選択する単語選択工程と、上記単語選択工程において選択された第2言語の単語を上記第1言語検索式のキーワードと置き換えて第2言語検索式とする置き換え工程と、からなることを特徴とする第1言語検索式を第2言語検索式に変換するための検索式変換方法。

【0026】この第1番目の発明の解決手段には、次の2つの検索式変換方法が含まれる。すなわち、第1番目の発明の検索式変換方法において、上記対訳工程と上記単語選択工程との間には、複数の対訳単語の順位付けを行う順位付け工程が挿入されていることを特徴とする検

索式変換方法。

【0027】上記の解決手段には、更に、次の検索式変換方法が含まれる。すなわち、上記の検索式変換方法において、上記順位付け工程における順位付けは、第2言語の単語共起頻度テーブルを用いて行われることを特徴とする検索式変換方法。

【0028】「第2番目の発明の解決手段」第1番目の発明の解決手段又はこれに含まれる上記2つの発明の解決手段のいずれかの発明の検索式変換方法をコンピュータに実行させるためのプログラムが記録された記録媒体。

【0029】「第3番目の発明の解決手段」第1の言語に基づいて構築された第1のキーワード特許文献検索システムと第2の言語に基づいて構築された第2のキーワード特許文献検索システムとの間に介在して両者との間で通信可能に接続されるための複数言語間の特許文献検索のためのインターフェースシステムであって、この複数言語間の特許文献検索のためのインターフェースシステムは、上記第1のキーワード特許文献検索システムのための第1言語検索式を入力するための第1検索式入力手段と、上記第1言語検索式により上記第1のキーワード特許文献検索システムにおいて検索されることにより適合した特許文献にかかる特許文献情報に基づいて、上記適合した特許文献の中から上記第2の言語によって記載されたパテントファミリー文献番号を抽出するファミリー番号抽出手段と、上記第1言語検索式のキーワードのそれぞれを単数又は複数の第2の言語の単語に対訳する対訳手段と、上記対訳手段によって対訳された第2言語の単数又は複数の単語の中から、上記ファミリー番号抽出手段で抽出された番号を有する第2言語のファミリー文献に付与されている第2言語キーワードに一致する第2言語の単語を選択する単語選択手段と、上記単語選択手段によって選択された第2言語の単語を上記第1言語検索式のキーワードと置き換えて第2言語検索式とする置き換え手段と、を備えており、上記第2言語検索式により、上記第2の言語によって記載された特許文献を対象にして、第2のキーワード特許文献検索システムに、キーワード検索を実行させるものであることを特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【0030】この第3番目の発明の解決手段には、次の2つのインターフェースシステムが含まれる。すなわち、第3番目の発明の複数言語間の特許文献検索のためのインターフェースシステムにおいて、このインターフェースシステムは、更に、上記対訳手段によって対訳された第2言語の複数の対訳単語に対して順位付けを行う順位付け手段を備えていることを特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【0031】上記の解決手段には、更に、次のインターフェースシステムが含まれる。すなわち、上記の発明の

9
複数言語間の特許文献検索のためのインターフェースシステムにおいて、上記順位付け手段により行われる順位付けは、第2言語の単語共起頻度テーブルを用いて行われることを特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【0032】「第4番目の発明の解決手段」第3番目の発明の解決手段又はこれに含まれる上記2つの発明の解決手段のいずれかの発明の複数言語間の特許文献検索のためのインターフェースシステムにおいて、このインターフェースシステムは、更に、上記置き換え手段によって、上記第1言語検索式のキーワードの全てが、上記第2言語のファミリー文献に付与されている第2言語キーワードに置き換えられなかったとき、第2の言語の特許文献の文面又は第2の言語のキーワードを表示するための第2言語文献表示手段と、上記第2言語文献表示手段によって表示された、第2の言語の特許文献の文面又は第2の言語のキーワード、から選んで重要なキーワードを入力又は選択するためのキーワード選択手段と、を備えており、上記キーワード選択手段によって入力又は選択されたキーワードにより、上記第2の言語によって記載された特許文献を対象にして、第2のキーワード特許文献検索システムに、キーワード検索を実行させるものであることを特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【0033】「第5番目の発明の解決手段」第3番目又は第4番目の発明の複数言語間の特許文献検索のためのインターフェースシステムにおいて、このインターフェースシステムは、更に、上記第2の言語によって記載された特許文献を対象にして、第2のキーワード特許文献検索システムによって実行されたキーワード検索の結果を出力するための出力手段を備えること、を特徴とする複数言語間の特許文献検索のためのインターフェースシステム。

【0034】「第6番目の発明の解決手段」第3番目又は第4番目の発明の複数言語間の特許文献検索のためのインターフェースシステムをコンピュータシステムに実現させるためのプログラムが記録された記録媒体。

【0035】「第7番目の発明の解決手段」それぞれが異なる言語で構築された複数の文献検索システムを含み、同一の内容の文献が上記各文献検索システムのデータベースに散在している複数言語検索システムのための検索方法であって、この検索方法は、一つの言語により作成された質問を上記複数の言語のうちのこの言語以外の言語に対応する翻訳候補を作成するための対訳辞書と、上記一つの言語により作成された質問に対する回答文献の一部に存在し、各言語により記述された同一の内容の文献とを用いて、上記翻訳候補の中から当該質問に対応する各言語の質問を選定することを特徴とする複数言語検索システムのための検索方法。

【0036】「第8番目の発明の解決手段」それぞれが

異なる言語で構築された複数の文献検索システムを含み、同一の内容の文献が上記各文献検索システムのデータベースに散在している複数言語検索システムであって、この複数言語検索システムは、一つの言語により作成された質問を上記複数の言語のうちのこの言語以外の言語に対応する翻訳候補を作成するための対訳辞書と、上記一つの言語により作成された質問に対する回答文献の一部に存在し、各言語により記述された同一の内容の文献とを用いて、上記翻訳候補の中から当該質問に対応する各言語の質問を選定するための質問言語変換手段を備えていることを特徴とする複数言語検索システム。

【0037】「第9番目の発明の解決手段」それぞれが異なる言語で構築された複数の文献検索システムを含み、同一の内容の文献が上記各文献検索システムのデータベースに散在している複数言語検索システムのためのプログラムが記録された記録媒体であって、この記録媒体に記録されたプログラムは、一つの言語により作成された質問を上記複数の言語のうちのこの言語以外の言語に対応する翻訳候補を作成するための対訳辞書と、上記一つの言語により作成された質問に対する回答文献の一部に存在し、各言語により記述された同一の内容の文献とを用いて、上記翻訳候補の中から当該質問に対応する各言語の質問を選定するための質問言語変換機能を上記複数言語検索システムに実現させるプログラムであることを特徴とする記録媒体。

【0038】「第10番目の発明の解決手段」第1の言語に基づいて構築された第1のキーワード特許文献検索システムと、第2の言語に基づいて構築された第2のキーワード特許文献検索システムと、上記2つのキーワード特許文献検索システムの間を介して両者との間で通信可能に接続されるための複数言語間の特許文献検索のためのインターフェースシステムとを備えた複数言語間の特許文献検索システムであって、上記インターフェースシステムは、上記第1のキーワード特許文献検索システムのための第1言語検索式を入力するための第1検索式入力手段と、上記第1言語検索式により上記第1のキーワード特許文献検索システムにおいて検索されることにより適合した特許文献にかかる特許文献情報に基づいて、上記適合した特許文献の中から上記第2の言語によって記載されたパテントファミリー文献番号を抽出するファミリー番号抽出手段と、上記第1言語検索式のキーワードのそれぞれを単数又は複数の第2の言語の単語に対訳する対訳手段と、上記対訳手段によって対訳された第2言語の単数又は複数の単語の中から、上記ファミリー番号抽出手段で抽出された番号を有する第2言語のファミリー文献に付与されている第2言語キーワードに一致する第2言語の単語を選択する単語選択手段と、上記単語選択手段によって選択された第2言語の単語を上記第1言語検索式のキーワードと置き換えて第2言語検索式とする置き換え手段と、を備えており、上記第2言語

検索式により、上記第2の言語によって記載された特許文献を対象にして、第2のキーワード特許文献検索システムに、キーワード検索を実行させるインターフェースシステムであることを特徴とする複数言語間の特許文献検索システム。

【0039】上記第1番目は第1言語検索式を第2言語検索式に変換するための検索式変換方法であり、上記第3番目から第5番目までの発明の複数言語間のインターフェースシステムは、第1のキーワード特許文献検索システムと第2のキーワード特許文献検索システムと協働して上記課題を解決するものである。また、第7番目及び第8番目の発明は、それぞれ、複数言語検索システムのための検索方法、及び、複数言語検索システムであつて、上記課題を解決するものである。

【0040】また、第10番目の発明は、第1言語キーワード特許文献検索システムと、第2言語キーワード特許文献検索システムと、2つのキーワード特許文献検索システムの間に介在するインターフェースシステムとをサブシステムとして包含する複数言語間の特許文献検索システムである。

【0041】更に、第2番目、第6番目及び第9番目の発明はこれらの発明をコンピュータにおいて実行させる乃至実現させるためのプログラムが記録された記録媒体に関する発明である。

【0042】ここで、これらの発明の複数言語間のインターフェースシステムは、第1のキーワード特許文献検索システム、及び、第2のキーワード特許文献検索システムと、商用回線又は専用回線を介して結ばれていることも可能であり、また、単一システムをなすコンピュータの内部で結ばれていることも可能である。それ故、更に、「通信可能に接続される」ということは、同一のコンピュータシステム内におけるデータのやりとりも、他のコンピュータシステムを結ぶ商用又は専用の通信回線を介して行われる通信も含むものである。

【0043】また、これまでに記載の発明において、第1の言語と第2の言語との関係は、それぞれ、本明細書の「発明の実施の形態」の欄中では、「日本語」と「英語」の関係あるいは日本特許文献検索システムと米国特許文献検索システムとして説明されており、このことから、他の用語についても相応の置き換え理解が必要とされる。また、「第1のキーワード特許文献検索システム」及び「第2のキーワード特許文献検索システム」は、それぞれ「パトリス、PATOLIS」（登録商標、本出願人、財団法人 日本特許情報機構）、及び、「IPS for DB 2」（IBM Intellectual Property Server for DB2, International Business Machine Corporation.）としてサービスされている検索システムを使用可能であるが、これらは互いに異なる複数の他の言語のシステムの関係であつてもよく、また、現にサービスされている、あるいは、今後サービスが開始される

他のキーワード特許文献検索システムであつてもよい。

【0044】

【発明の実施の形態】以下に、複数言語の特許文献のための機械検索システム等の各請求項に記載された発明を、日本語特許文献と米国特許文献を対象とする日米特許文献検索システムによって説明する。

【0045】図1は、本発明を大まかに説明するため一例として示した概念図、図2、図4は、図1とは異なる観点から本発明を説明するための概念図である。また、図3は、三つのサブシステム間の動作・情報の授受を時間をおって説明するためのチャートである。

【0046】日米特許文献検索システム1は、日本語キーワード特許文献検索システム（以下、日本特許文献検索システム2という。）、英語キーワード特許文献検索システム（以下、米国特許文献検索システム3という。）、及び、複数言語間の特許文献検索のためのインターフェースシステム（以下、単に、インターフェースシステムという。）4を備えている。日本特許文献検索システム2とインターフェースシステム4は通信線で結合されており、更にこのインターフェースシステム4と米国特許文献検索システム3とは商用通信回線5に接続されている。商用通信回線5にはユーザーのコンピュータすなわちクライアント6が任意に接続される。したがって、これら各システム及びクライアント6とは、通信可能になっている。

【0047】日本特許文献検索システム2、及び、米国特許文献検索システム3には、それぞれ上に例示したPATOLIS及びIPS for DB 2を利用することができる。

【0048】この実施例の動作は以下の通りである。すなわち、インターフェースシステム4において、サーチャーは日本語検索式を作成し、日本特許文献検索システム2に検索を実行させる。ここで、J-DB（図3）は日本語特許文献データベースを示している。検索結果をみて日本語検索式、キーワード（J-KW、図3）を変えながら、満足できる回答を得るまで試行錯誤を繰り返す。

【0049】満足できる日本語特許文献が得られた（ヒットした）とサーチャーが判断したとき、この日本語特許文献を指定するとともに、満足できる日本語特許文献がヒットしたことをインターフェースシステム4に入力するとファミリー番号抽出指令が出される。ファミリー番号抽出指令はヒットした日本語特許文献群の中から米国特許にファミリーを持つ特許を抽出しその米国特許番号を取り出すことを指示するものである。

【0050】なお、相互にファミリー関係にある日米特許番号は、その都度一方の特許文献に記載された優先権番号をタグにして他方の特許文献番号を求めるようにすることもできるし、予めこれらの関係を求めることによつて作成した日米特許番号対照テーブルを参照して求め

ることもできる。また、米国特許にファミリーを持つ特許が存在しない場合もあり、このときはPAJ (Patent Abstract of Japan) を利用することができる。

【0051】上記日本特許文献検索システム2における検索の検索式中に使用されている日本語キーワード(J-KW)の対訳英単語が日英対訳辞書から抽出される。これはインターフェースシステム4において行われる。

【0052】一つの日本語キーワード(J-KW)の対訳英単語は通常複数個存在するから、日本語検索式が例えば4つの日本語キーワードからなるとすれば、10から20の対訳英単語が抽出されることになる。

【0053】次に、インターフェースシステム4の備えている英単語共起頻度テーブル(共起TBL、図2)が参照され、上記の抽出された対訳英単語の中から共起性の高い英単語の組み合わせ候補が幾組か選出される。選出された共起性の高い英単語の組み合わせ候補に、先に指定された日本語特許文献に対応する米国特許文献に付されている英語キーワード(E-KW)があるか否かが検査される。組み合わせ候補にこのような英語キーワードが無いときは、つまり日本語検索式の要素を完全に英語キーワードで埋め合わせることができないときは、機械的(自動的)に英語検索式を作成することができないので、人手による検索式作成作業に入る。

【0054】上記検査により英語キーワードがあるときは、この英語キーワード(E-KW)が英語検索式のキーワードとして確定(複数ある候補の中の 하나가英語検索式のキーワードとして決定される)され、この英語検索式が米国特許文献検索システム3に送信され、米国特許文献検索システム3において米国特許文献データベース(E-DB)に対して送信された英語検索式によって検索が実行される。なお、このときの上記検査の結果に基づいて日英対訳辞書の更新を行うことができる。

【0055】検索結果が満足できるものであれば、回答の出力に進む。一方、不満足るときには先にファミリー番号抽出指令によって取り出された米国特許文献番号をもとに米国特許文献検索システム3において、番号検索が実行され、米国特許文献の明細書、アブストラクト、特許分類、この文献に付されている英語キーワード等の英語キーワード選出に有効なデータがインターフェースシステム4に取り込まれる。

【0056】インターフェースシステム4に取り込まれたデータは、サーチャーのディスプレイ等に表示あるいは印刷され、サーチャーはこれを参考にして英語検索式を作成する。

【0057】サーチャーによって作成された英語検索式は米国特許文献検索システム3に送信され、これを受信した米国特許文献検索システム3は米国特許文献データベース(E-DB)によって検索を実行する。英語検索式の作成又は修正が何度か行われることにより、目的の米国特許文献に到達し、この文献が回答出力される。な

お、この場合にも日英対訳辞書の更新が行われるようにすることが好ましい。

【0058】次に、検索式の自動作成の過程を示す具体的な検索実験例、熱変位による加工誤差を補正する工作機械に関する特許文献をサーチした例、を用いて本発明を説明する。

【0059】この場合日本語キーワードとして、工作機械、熱変位、位置補正(又は、補償)が、また、日本語検索式は

工作機械 * 熱変位 * (位置補正 + 補償) とすることが考えられる。

【0060】この日本語検索式を日本特許文献検索システム2に入力、検索実行をすると、全67件の日本語特許文献の回答が得られた。この回答から米国にファミリーを有する文献を抽出すると7件が抽出された。この出力を受けて米国特許文献検索システム3からファミリーの米国特許文献の諸データを出力する。なお、記載されている発明の内容の凝集度を考慮すると、出力内容はタイトル、アブストラクト程度が適当であろう。

【0061】一方、前記日本語検索式は

「工作、機械、工作機械、熱、変位、熱変位、(位置、補正、位置補正) + 補償)」

に分ち書きされ、日英対訳辞書により、

工作 workmanship, construct, building

機械 machine, machinery

工作機械 machine tool

熱 heat, fever

変位 displacement

熱変位 thermal displacement

位置 position, location

補正 revision, correction, compensation

補償 compensation, reparation

に翻訳される。これらの訳語の中で、machine tool や thermal displacement のような複合語として存在するのはそのまま複合語が採用され、position, location, revision, correction, compensation 及び (machine tool) (thermal displacement) を単語共起頻度テーブル(米国特許文献に関するデータベースにおいて共に出現する頻度が高い組み合わせを頻度順にしたテーブル)と突き合わせると、(machine tool) (thermal displacement) correction, position, compensation は共に出現頻度が高い組み合わせ候補となる。次に日本語キーワードにおいて複合語として取り扱った(位置補正)は英語における接続処理を施し、correcting position となる。

【0062】次に、以上のキーワード(machine tool)、(thermal displacement)、(correcting position) 及び compensation がファミリーの米国特許文献のデータとして用いられているかどうかを検査、確認し、用いられている場合には、日本語キーワードがこれ

らによって置き換えられ日本語検索式の式（論理構造）と同じ構造を持つ英語検索式、つまり、
 ((machine tool) and (thermal displacement) and
 (compensation or (correcting position))
 を作成する。

【0063】この英語検索式が米国特許文献検索システム3に入力され19件の米国特許文献が回答として得られ、出力された。

【0064】以上は、英単語組み合わせ候補がファミリーの米国特許文献のデータで使用されている場合である。ところが常にこのようにできるとは限らず、英単語組み合わせ候補の一部又は全部が使用されていないような場合、つまり、日本語検索式のキーワードを完全に埋め合わせることができないような場合には、自動的に英語検索式を作成することができないので、ファミリーの米国特許文献のタイトル、アブストラクト等を形態素解析等の処理をして、抽出した英語キーワードと共に英単語組み合わせ候補（接続処理済み）を表示させ、サーチャーはこれらを参考に英文検索式を作成することになる。このとき、ファミリーの米国特許文献に付された英語キーワード及び米国特許分類あるいはIPCを表示して英語検索式の作成に利用することも可能であり、このようにすることにより英語検索式の作成が容易となる。

【0065】

【発明の効果】本発明では、特許文献の Patent ファミ

リー関係が積極的に利用され、日本語で構築された特許文献データベースを検索して得られた成果、及び、このときの検索式が英語で構築された特許文献データベースの検索に有効に利用されるので、日本語検索式を使用して日本語特許文献を入手すると同時に、英語検索式が自動的に作成されて米国特許文献も入手できることができるという効果を奏する。また、自動的に検索式が作成されないときでも、人による英語検索式の作成が容易になるという効果を奏する。

【図面の簡単な説明】

【図1】本発明を大まかに説明するための一例として示す概念図である。

【図2】図1とは異なる観点から本発明を説明するための概念図である。

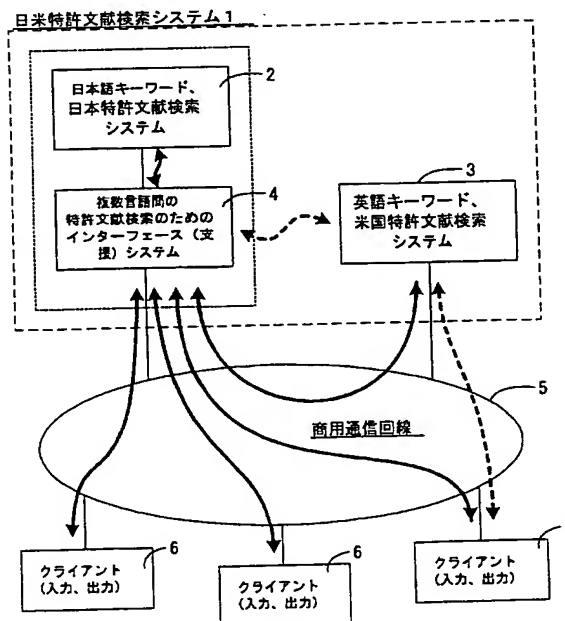
【図3】三つのサブシステム間の動作・情報の授受を時間をおって説明するためのチャートである。

【図4】図1とは異なる観点から本発明を説明するための概念図である。

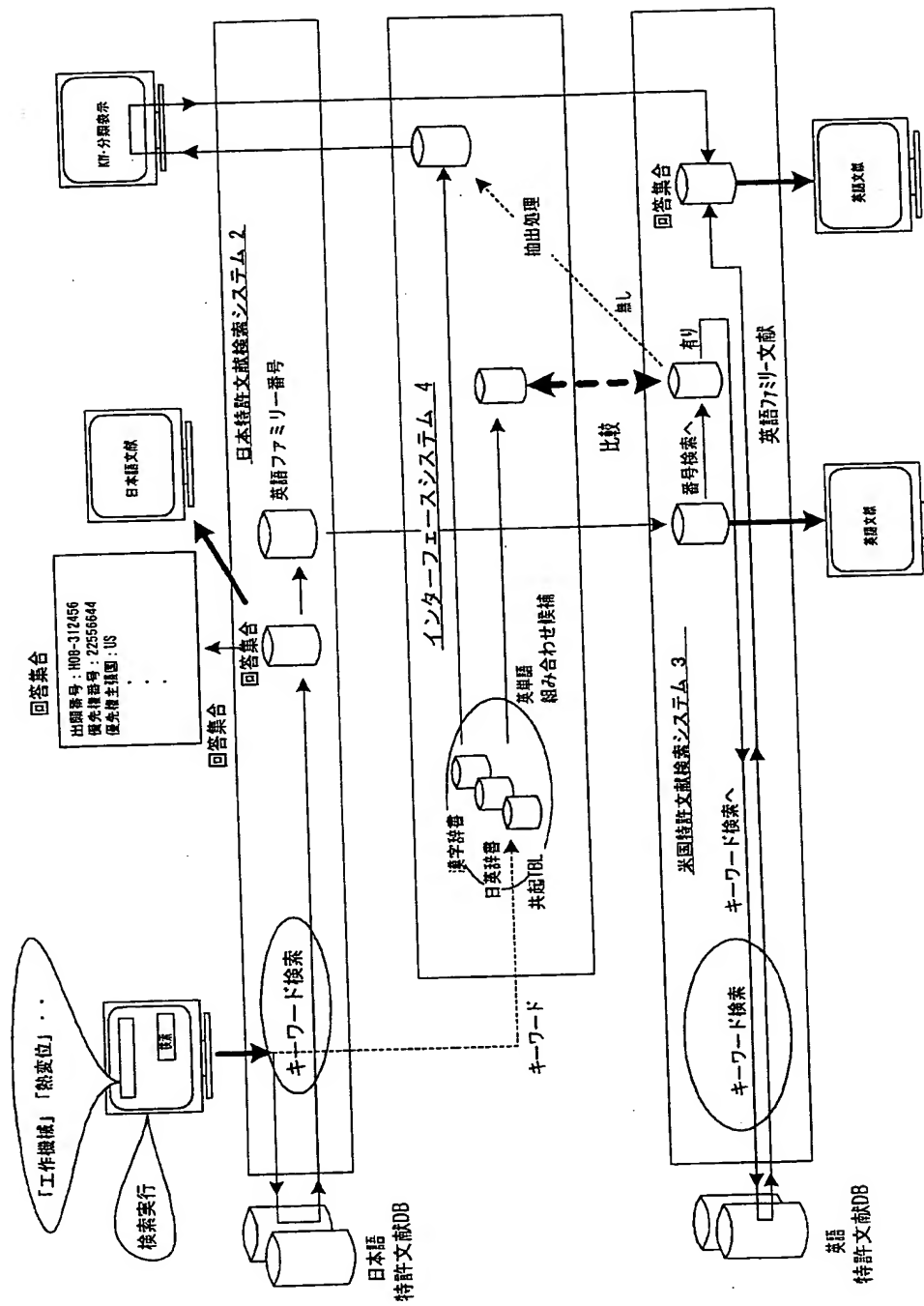
【符号の説明】

- 1 日米特許文献検索システム
- 2 日本語特許文献検索システム
- 3 米国特許文献検索システム
- 4 インターフェースシステム
- 5 商用通信回線
- 6 クライアント

【図1】

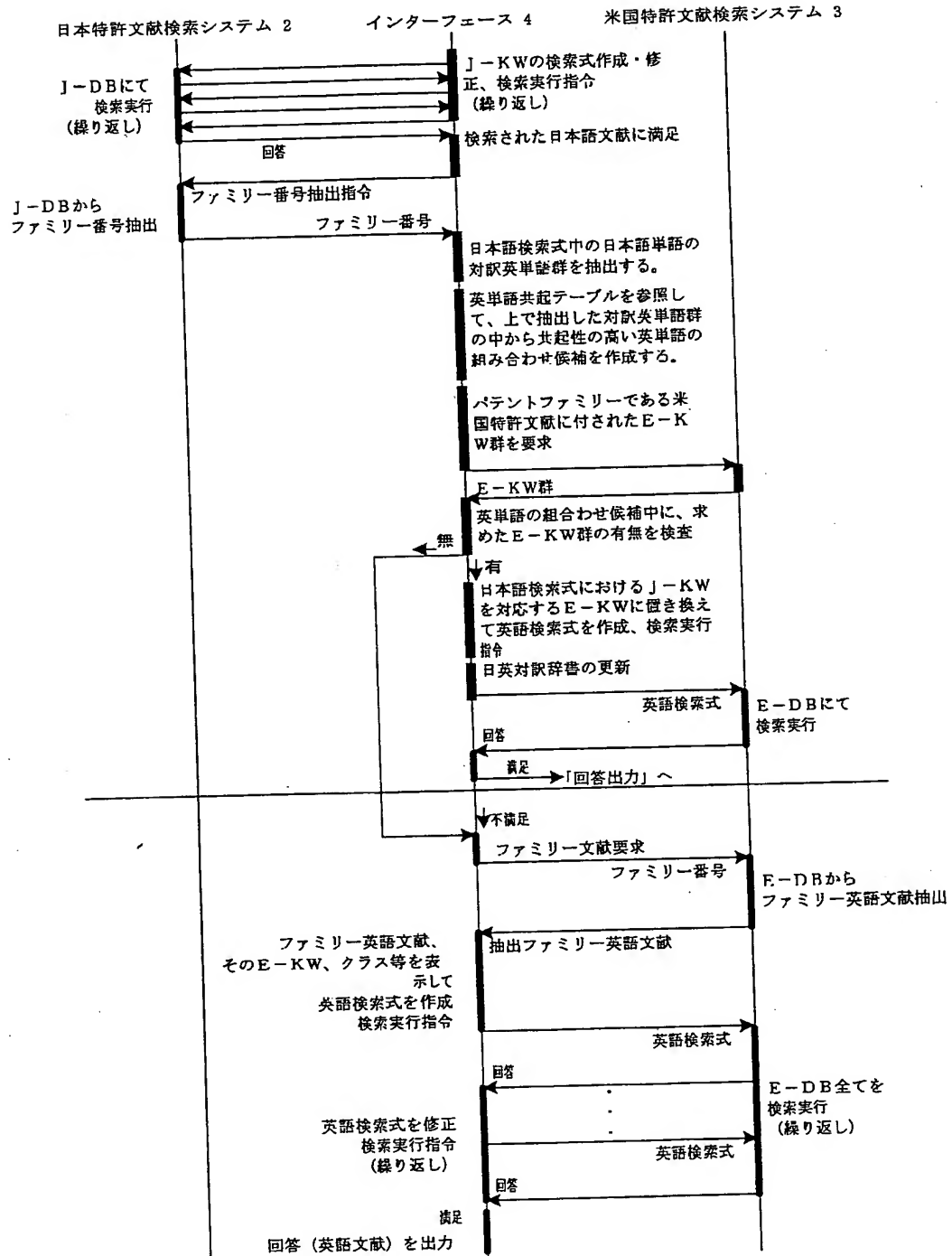


【図2】

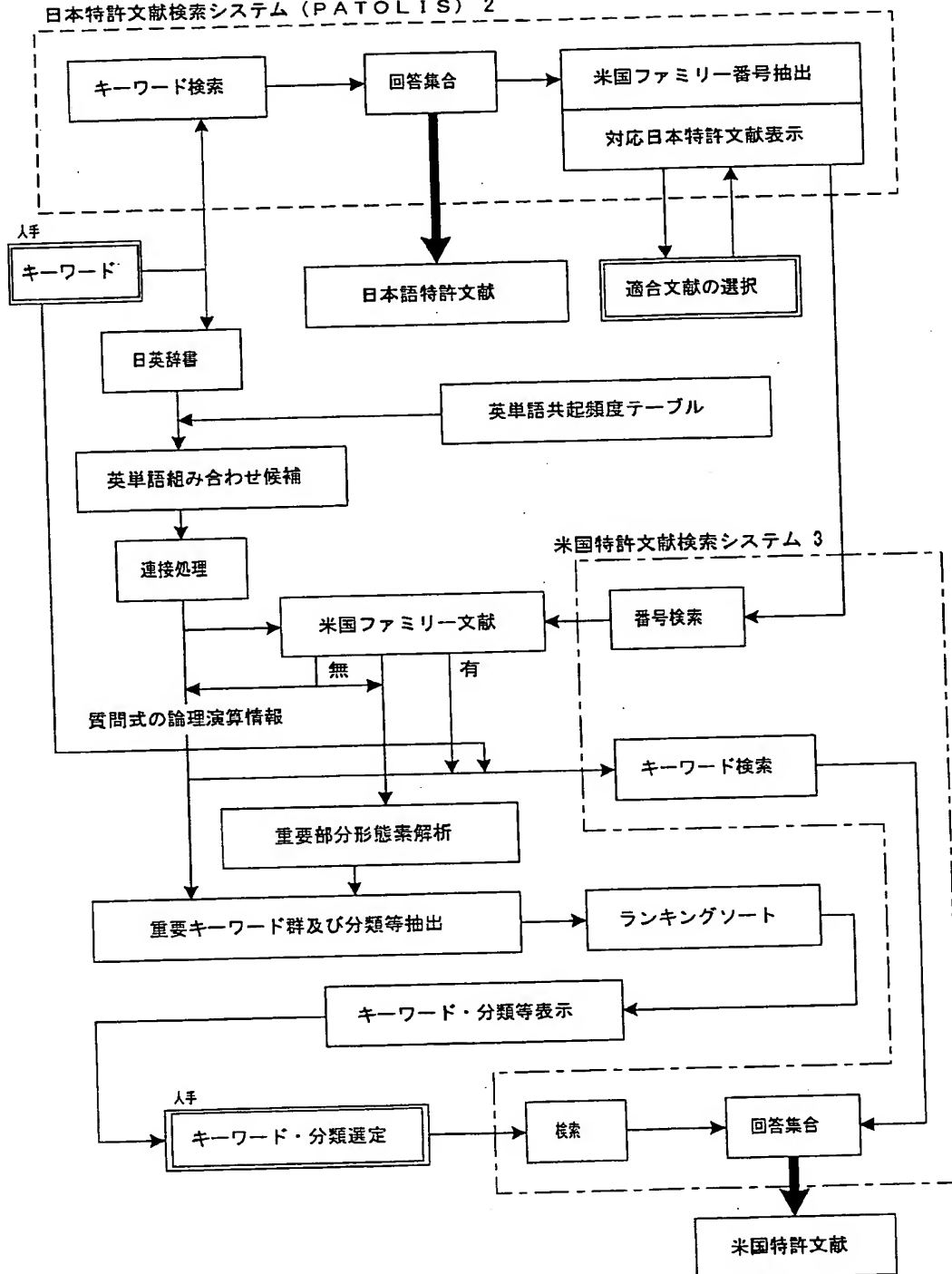


【図3】

複数言語間の特許文献検索システム



日本特許文献検索システム (PATOLIS) 2



フロントページの続き

(72)発明者 樋口 重人

東京都江東区東陽 4 丁目 1 番 7 号 佐藤ダ
イヤビルディング 財団法人 日本特許情
報機構内

(72)発明者 福井 雅敏

東京都江東区東陽 4 丁目 1 番 7 号 佐藤ダ
イヤビルディング 財団法人 日本特許情
報機構内

F ターム(参考) 5B075 NK02 PP23 PP25 PP26 PQ02

PR04 QP03

5B091 AA03 EA24

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.